

Statistical Mechanics Analysis of the Continuous Number Partitioning Problem

F. F. Ferreira* and J. F. Fontanari
 Instituto de Física de São Carlos
 Universidade de São Paulo
 Caixa Postal 369
 13560-970 São Carlos SP
 Brazil

Abstract

The number partitioning problem consists of partitioning a sequence of positive numbers $\{a_1, a_2, \dots, a_N\}$ into two disjoint sets, \mathcal{A} and \mathcal{B} , such that the absolute value of the difference of the sums of a_j over the two sets is minimized. We use statistical mechanics tools to study analytically the Linear Programming relaxation of this NP-complete integer programming. In particular, we calculate the probability distribution of the difference between the cardinalities of \mathcal{A} and \mathcal{B} and show that this difference is not self-averaging.

PACS: 87.10.+e; 64.60.Cn

Keywords: Number partitioning, Linear Programming relaxation

1 Introduction

Although most of the statistical mechanics analyses of stochastic versions of combinatorial optimization problems have focused mainly on the calculation of the average cost of the global optima [1], the tools of equilibrium statistical mechanics can also be used to evaluate the average performance of simple heuristics as well as that of relaxed versions of the original problem

*Corresponding author. E-mail: fagundes@ifsc.sc.usp.br

[2]. In this paper we study both numerically and analytically the Linear Programming (LP) relaxation of a classical NP-complete integer programming problem, namely, the number partitioning problem [3, 4].

The number partitioning problem (NPP) is stated as follows [4]. Given a sequence of positive numbers $\{a_1, a_2, \dots, a_N\}$, the NPP consists of partitioning them into two disjoint sets \mathcal{A} and \mathcal{B} such that the difference

$$\left| \sum_{a_j \in \mathcal{A}} a_j - \sum_{a_j \in \mathcal{B}} a_j \right| \quad (1)$$

is minimized. Alternatively, we can search for the Ising spin configurations $\mathbf{s} = (s_1, \dots, s_N)$ that minimize the energy or cost function

$$E(\mathbf{s}) = \left| \sum_{j=1}^N a_j s_j \right|, \quad (2)$$

where $s_j = 1$ if $a_j \in \mathcal{A}$ and $s_j = -1$ if $a_j \in \mathcal{B}$. Also of interest is the problem of constrained partitions, in which the difference between the cardinalities of sets \mathcal{A} and \mathcal{B} is fixed, i.e.,

$$m = \frac{1}{N} \left| \sum_{j=1}^N s_j \right|, \quad (3)$$

so that the cardinality of the largest set is $N(1+m)/2$. Henceforth we will restrict our analysis to the case where the a_j 's are statistically independent random variables uniformly distributed in the unit interval.

The interest in the NPP stems mainly from the remarkable failure of the stochastic heuristic simulated annealing to find good solutions to it, as compared with the solutions found by deterministic heuristics [5]. The reason for that failure is probably due to the existence of order of 2^N local minima whose energies are of order of $1/N$ [8], which undermines the usual strategy of exploring the space of configurations $\{\mathbf{s}\}$ through single spin flips. It is interesting to note that a very simple deterministic heuristic, the differencing method of Karmarkar and Karp, can find with high probability solutions whose energies are of order of $1/N^{\alpha \log N}$ for some $\alpha > 0$ [6, 7]. For large N , however, the energies of the solutions found by the differencing method are orders of magnitude higher than those predicted by theoretical analyses, which indicate that the average global optimal energy \bar{E}_0 is of order of $\sqrt{N} 2^{-N}$ for unconstrained partitions [4, 8]. A recent exact calculation of this quantity yielded $\bar{E}_0 = \sqrt{2\pi N/3} 2^{-N}$ [9]. It must be noted that, in contrast with combinatorial problems for which the global optimal energy is

extensive [1], for the NPP this energy is not self-averaging [8, 9] and hence \bar{E}_0 cannot be viewed as a realization independent minimal energy.

In the LP relaxation we relax the integrality requirement on the Ising variables s_i so that they become real variables, i.e., $s_i \in (-\infty, \infty)$. In order to keep these variables finite we impose a spherical constraint on the norm of the solutions,

$$\sum_i^N s_i^2 = N. \quad (4)$$

Obviously, minimizing the square of the cost (2) with s_i real but constrained to obey the condition (4) yields a lower bound to the optimal (square) cost of the corresponding integer programming problem. Moreover, a simple gradient descent dynamics suffices to attain those bounds numerically. In fact, using a Lagrangian multiplier to handle the constraint (4) we find that the following dynamics minimizes the NPP energy,

$$\frac{\partial s_i}{\partial t} = -\eta \zeta \left(a_i - \frac{\zeta}{N} s_i \right) \quad i = 1, \dots, N \quad (5)$$

where $\zeta = \sum_j a_j s_j$ and η is an arbitrarily small parameter that determines the step-size of the descent.

The remainder of this paper is organized as follows. In section 2 we show that the LP relaxation yields a trivial lower bound (i.e. the LP cost is zero) for unconstrained partitions. That analysis yields, nonetheless, some interesting pieces of information as, for instance, the average energy \bar{E}_c obtained by clipping (i.e. taking the sign of) the spins of the global optimal configurations of the LP relaxation. The average performance of the clipping heuristic is studied in section 3, where it is shown that \bar{E}_c tends to the average energy of a randomly chosen Ising spin configuration, $\bar{E}_c \rightarrow \sqrt{2N/3\pi}$. In section 4 we calculate the probability distribution of the difference between the cardinalities of \mathcal{A} and \mathcal{B} for the LP global optimal configurations, $\mathcal{P}_c(m)$, and show that m is not self-averaging. Finally, in section 5 we present some concluding remarks.

2 Linear Programming relaxation

In the canonical ensemble formalism of the statistical mechanics the average value of the optimal energy for unconstrained partitions is given by

$$\bar{E}_u = - \lim_{T \rightarrow 0} T \langle \ln Z_u \rangle_a, \quad (6)$$

where $Z_u(T)$ is the partition function

$$Z_u(T) = \prod_i \int_{-\infty}^{\infty} ds_i \delta \left(N - \sum_j s_j^2 \right) \exp \left[-\frac{E(\mathbf{s})}{T} \right] \quad (7)$$

with $E(\mathbf{s})$ given by Eq. (2). Here $\delta(x)$ is the Dirac delta and T is the temperature. The notation $\langle \dots \rangle_a$ stands for the average over the random variables a_i . The limit $T \rightarrow 0$ in Eq. (6) ensures that only the states that minimize $E(\mathbf{s})$ will contribute to Z_u . We now proceed with the explicit evaluation of the partition function (7). Using the integral representation of the Dirac delta function we write

$$\begin{aligned} Z_u(T) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{d\zeta d\tilde{\zeta}}{2\pi} \int_{-\infty}^{\infty} \frac{dx}{2\pi} e^{i\zeta\tilde{\zeta} + ixN - |\zeta|/T} \\ &\quad \times \prod_j \int_{-\infty}^{\infty} ds_j \exp \left[-ixs_j^2 - is_j a_j \tilde{\zeta} \right]. \end{aligned} \quad (8)$$

The integrals over s_j and $\tilde{\zeta}$ can easily be performed yielding

$$\begin{aligned} Z_u(T) &= \sqrt{\frac{\pi^{N-3}}{4\mathcal{M}_2}} \int_{-\infty}^{\infty} dx e^{ixN} (ix)^{(1-N)/2} \\ &\quad \times \int_{-\infty}^{\infty} d\zeta e^{-|\zeta|/T - ix\zeta^2/\mathcal{M}_2} \end{aligned} \quad (9)$$

where $\mathcal{M}_2 = \sum_j a_j^2$. At this stage the integral over ζ can be readily carried out by assuming $T \rightarrow 0$. The result is simply

$$Z_u(T) = T \sqrt{\frac{\pi^{N-3}}{4\mathcal{M}_2}} \int_{-\infty}^{\infty} dx (ix)^{1/2} e^{NG_u(x)} \quad (10)$$

where

$$G_u(x) = ix + -\frac{1}{2} \ln(ix). \quad (11)$$

In the limit of large N , the integral over x can be evaluated using the saddle-point method. Noting that the saddle-point is the imaginary $x_s = 1/2i$, the unconstrained partition function is finally written as

$$Z_u(T) = T \sqrt{\frac{1}{2\pi^2 N \mathcal{M}_2}} (1 + \ln 2\pi)^{N/2}. \quad (12)$$

Since Z_u decreases linearly with decreasing T , Eq. (6) yields $\bar{E}_u = 0$. This result was verified numerically using the gradient descent dynamics (5) together with an adaptive prescription to decrease the step-size η during the descent.

3 Clipping heuristic

An easy-to-implement procedure to generate Ising solutions from the LP relaxation solutions is to take the sign of the relaxed spins. The average cost associated to this clipping procedure is given by

$$\bar{E}_c = \lim_{T \rightarrow 0} \left\langle \left\langle \left| \sum_j a_j \text{sign}(s_j) \right| \right\rangle_T \right\rangle_a \quad (13)$$

where $\langle \dots \rangle_T$ stands for a thermal average taken with the Gibbs probability distribution, i.e., $\exp[-E(\mathbf{s})/T]/Z_u$. The zero-temperature limit ensures that only configurations that minimize the relaxed cost (2) will contribute to this average. To evaluate Eq. (13) we introduce the auxiliary energy

$$E_{clip}(\mathbf{s}) = E(\mathbf{s}) + h \left| \sum_j a_j \text{sign}(s_j) \right| \quad (14)$$

with $E(\mathbf{s})$ given by Eq. (2). Hence

$$\bar{E}_c = - \lim_{T \rightarrow 0} T \left. \frac{\partial \langle \ln Z_{clip} \rangle_a}{\partial h} \right|_{h=0} \quad (15)$$

where Z_{clip} is the partition function (7) with E replaced by E_{clip} . Introducing the auxiliary parameter $v = \sum_j a_j \text{sign}(s_j)$ through a Dirac delta function, the calculation of Z_{clip} becomes analogous to that presented before, and so we will present the final results only. We find

$$\begin{aligned} \bar{E}_c &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dv |v| \int_{-\infty}^{\infty} d\tilde{v} e^{iv\tilde{v}} \left\langle \prod_j \cos(\tilde{v} a_j) \right\rangle_a \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dv |v| \int_{-\infty}^{\infty} d\tilde{v} e^{iv\tilde{v}} \left(\frac{\sin \tilde{v}}{\tilde{v}} \right)^N. \end{aligned} \quad (16)$$

Assuming that \bar{E}_c does not increase linearly with N , in the thermodynamic limit only the regions close to the origin ($\tilde{v} = 0$) will contribute to the integral over \tilde{v} in Eq. (16). Hence using $\sin \tilde{v}/\tilde{v} \approx -\tilde{v}^2/6$ yields

$$\bar{E}_c = \sqrt{\frac{2N}{3\pi}}. \quad (17)$$

We have found a remarkably good agreement between this theoretical prediction and the properly averaged cost obtained by clipping the spherical

spins in LP minima generated by the gradient descent dynamics (5). Interestingly, for large N the cost (17) is identical to the average energy of a randomly chosen Ising configuration \mathbf{s} , defined by

$$\bar{E}_r = 2^{-N} \prod_i \int_0^1 da_i \sum_{s_i=\pm 1} \left| \sum_i a_i s_i \right|, \quad (18)$$

which thus demonstrate the complete failure of the clipping heuristic.

4 Probability distribution of cardinalities

As the distinct LP global minima will have, in general, different cardinalities, in this section we calculate analytically the probability distribution of the cardinalities difference defined by

$$\mathcal{P}_c(m) = \lim_{T \rightarrow 0} N \left\langle \left\langle \delta \left(Nm - \left| \sum_j s_j \right| \right) \right\rangle_T \right\rangle_a. \quad (19)$$

Using the definition of the thermal average this equation is rewritten as

$$\mathcal{P}_c(m) = \lim_{T \rightarrow 0} N \left\langle \frac{Z_m}{Z_u} \right\rangle_a, \quad (20)$$

where

$$Z_m = \prod_i \int_{-\infty}^{\infty} ds_i \delta \left(N - \sum_j s_j^2 \right) \exp \left[-\frac{E(\mathbf{s})}{T} \right] \delta \left(Nm - \left| \sum_j s_j \right| \right) \quad (21)$$

and Z_u is given by Eq. (7). The calculation of Z_m is a little more complicated than that of Z_u , as it involves the evaluation of an additional integral due to the extra delta function. However, since the steps are essentially the same in both calculations we will present the final result only. In the limits $T \rightarrow 0$ and $N \rightarrow \infty$ we find

$$Z_m = \frac{T}{N} \sqrt{\frac{1}{\pi^3 \mathcal{V}}} (1 + \ln 2\pi)^{N/2} \exp \left(-\frac{N \mathcal{M}_2}{2\mathcal{V}} m^2 \right), \quad (22)$$

where $\mathcal{V} = \sum_j a_j^2 - N \left(\frac{1}{N} \sum_j a_j \right)^2$. In the limit $N \rightarrow \infty$ we use the self-averaging property,

$$\frac{1}{N} \sum_i f(a_i) = \int_0^1 da f(a) \quad (23)$$

for any function f , to write $\mathcal{M}_2/N = 1/3$ and $\mathcal{V}/N = 1/12$ so that Eq. (20) becomes

$$\mathcal{P}_c(m) = \sqrt{\frac{8N}{\pi}} e^{-2Nm^2} \quad m \geq 0. \quad (24)$$

Hence, the mean is $\langle m \rangle = 1/\sqrt{2N\pi}$ and the variance, $\sigma_m^2 = (\pi - 2)/4\pi N$. An important quantity is the ratio $r_m = \sqrt{\sigma_m^2}/\langle m \rangle$, whose vanishing determines the self-averageness of the random variable m . In figures 1(a) and 1(b) we present the results of numerical experiments to estimate the dependence on N of $\langle m \rangle$ and r_m , respectively, for three types of configurations: (i) the global minima of the original NPP obtained through the exhaustive search in the Ising configuration space for $N \leq 26$; (ii) the legal, Ising configurations obtained with the differencing method (we refer the reader to ref. [7] for a clear presentation of this heuristic); and (iii) the global minima of the LP relaxation obtained with the dynamics (5). We note the very good agreement between the latter estimate and the analytical predictions. In all cases, the mean $\langle m \rangle$ decreases like $N^{-1/2}$ as N increases, while r_m tends to a nonzero value ($r_m = \sqrt{\pi/2 - 1} \approx 0.755$ for the LP relaxation), indicating that m is not self-averaging even in the large N limit, i.e., the values of m associated to the configurations under study depend on the specific realization of the set of random variables $\{a_j\}$.

5 Conclusion

In this paper we have illustrated the usefulness of equilibrium statistical mechanics tools to investigate analytically the average performance of standard relaxation procedures to generate lower bounds to integer programming problems, as well as to characterize specific properties of the minima. The failures of the LP relaxation and the clipping heuristic to produce relevant results for the NPP yield additional evidence to the extreme difficulty of devising heuristics to find near-optimal solutions to that problem.

Acknowledgments

We thank P. Moscato for useful discussions and suggestions. The work of JFF was supported in part by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). FFF is supported by FAPESP.

References

- [1] M. Mézard, G. Parisi and M. A. Virasoro, Spin Glass Theory and Beyond, World Scientific, Singapore, 1987.
- [2] J. Inoue, J. Phys. A 30 (1997) 1047.
- [3] M. R. Garey and D. S. Johnson, Computers and Intractability: A Guide to the Theory of NP-Completeness, Freeman, San Francisco, 1979.
- [4] N. Karmarkar, R. M. Karp, G. S. Lueker and A. M. Odlyzko, J. Appl. Prob. 23 (1986) 626.
- [5] D. S. Johnson, C. R. Aragon, L. A. McGeoch and C. Schevon C, Operations Research 39 (1991) 378.
- [6] B. Yakir, Math. Oper. Res. 21 (1996) 85.
- [7] M. F. Argüello, T. A. Feo and O. Goldschmidt, Computers Ops. Res. 23 (1996) 103.
- [8] F. F. Ferreira and J. F. Fontanari, J. Phys. A 31 (1998) 3417.
- [9] S. Mertens, cond-mat preprint/9807077.

Figure caption

Fig. 1 (a) Average cardinalities difference as a function of $1/N^{1/2}$ and (b) ratio between the standard deviation and the average cardinalities difference as a function of N . The convention is \bigcirc (LP relaxation), ∇ (exhaustive search) and $*$ (differencing method). The solid curves are the theoretical predictions for the LP relaxation.

